

# Tabla de Contenido

<b>Introducción</b>	<b>1</b>
<b>1. Marco Teórico</b>	<b>5</b>
1.1. Red neuronal Convolutacional (CNN)	5
1.1.1. Capa convolutacional	7
1.1.2. Capa de <i>Pooling</i>	8
1.1.3. Capa <i>Fully connected</i>	9
1.2. <i>Residual Network</i> (ResNet)	9
1.3. Detección de objetos	11
1.3.1. Faster R-CNN	11
1.3.2. Mask R-CNN	14
1.3.3. YOLOv3 ( <i>You Only Look Once</i> )	16
1.3.4. RetinaNet	22
1.3.5. TridentNet	25
1.4. Métricas de desempeño	27
1.5. Estado del arte	30
<b>2. Metodología</b>	<b>32</b>
2.1. Configuración experimental	32
2.1.1. Modelos de detección	32
2.1.2. <i>Frameworks</i> de entrenamiento	33
2.1.3. Entorno de pruebas	33
2.2. Conjunto de datos	34
2.2.1. DeepFashion2	34
2.2.2. ModaNet	36
2.3. Preparación de los datos	37
2.3.1. Darknet	38
2.3.2. Detectron y SimpleDet	38
2.4. Entrenamiento de los modelos	39
2.4.1. Darknet	39
2.4.2. Detectron	40
2.4.3. SimpleDet	41
2.5. Aplicación en <i>Retrieval</i>	42
<b>3. Resultados</b>	<b>44</b>
3.1. Tiempos de inferencia	44

3.2. Detección . . . . .	45
3.2.1. Conjunto de datos externo . . . . .	50
3.3. <i>Retrieval</i> . . . . .	53
<b>4. Análisis</b>	<b>55</b>
4.1. Deficiencias en métrica AP . . . . .	56
4.1.1. Retrieval . . . . .	59
<b>Conclusión</b>	<b>60</b>
<b>Bibliografía</b>	<b>62</b>